JOURNAL OF MULTIVARIATE ANALYSIS 34, 141–155 (1990)

Optimal Designs for Multivariate Interpolation

M. C. SPRUILL

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332

Communicated by the Editors

Optimal interpolation designs are given for estimation of the value of a linear functional of an unknown polynomial mean. For example, optimal designs are given for estimation of the average value of a multivariate polynomial over a sphere of radius $a<r$ when observations are available only in the region $r \leqslant \|x\| \leqslant R$.
© 1990 Academic Press, Inc.

1. INTRODUCTION

For all the cases covered in this paper it is assumed that a fixed number N of observations are to be taken, and that for any choice of points $\{x_1, ..., x_N\}$, not necessarily distinct, in the set X of available values of the design variable the uncorrelated random variables $\{Y(x_1), ..., Y(x_N)\}$ can be observed. Here

$$Y(x_i) = \theta(x_i) + \varepsilon_i, \qquad E(\varepsilon_i) = 0, \qquad E(\varepsilon_i^2) = \sigma^2,$$

where the mean θ is a polynomial and the x_i are d-dimensional vectors. The means θ which we treat are of two types. Either

$$\theta(x) = \sum_{\|\alpha\|_\infty \leqslant m-1} c_\alpha x^\alpha \tag{1.1}$$

or

$$\theta(x) = \sum_{\|\alpha\|_1 \leqslant m-1} c_\alpha x^\alpha, \tag{1.2}$$

where $x^\alpha = \prod_{i=1}^d x_i^{\alpha_i}$, $\|x\|_p$ denotes the usual p-norm of a vector $x \in R^d$, $1 \leqslant p \leqslant \infty$, and $\alpha_i \geqslant 0$ are integers.

Received January 24, 1990.
AMS 1980 subject classification: primary 62K05.
Key words and phrases: estimation, extremal polynomials, regression.

0047-259X/90 $3.00
Copyright © 1990 by Academic Press, Inc.
All rights of reproduction in any form reserved.

The unknown parameters c_α form a linear space of dimension m^d in the case of (1.1) or $\binom{m+d-1}{d}$ in the case of (1.2). In either case the quantities we seek optimal designs to estimate are linear functionals τ defined on this linear space. The designs reported here can be verified to be optimal for the minimum variance linear unbiased estimation of these quantities by using the following extension of Elfving's theorem (see Karlin and Studden [10]). In the theorem the space of means, whichever model is used, is denoted by Θ while τ is a linear functional defined on Θ. Let $\Delta = \{\theta \in \Theta : (\tau, \theta) = 1\}$ and Ξ be the finitely supported Borel probability measures on X.

THEOREM 1.1. *Suppose there is a point $\theta_0 \in \Delta$ and a design $\xi_0 \in \Xi$ satisfying*

(i) $\{x \in X: |\theta_0(x)| = \max_{z \in X} |\theta_0(z)|\} \supset S(\xi_0)$,

(ii) *for some non zero β and all $\theta \in \Theta$, $\int \theta(x)\, \theta_0(x)\, d\xi_0(x) = \beta(\tau, \theta)$,*

and

(iii) $\int \theta^2(x)\, d\xi_0(x) = 0$ *entails* $(\tau, \theta) = 0$. *Then ξ_0 is optimal for estimation of (τ, θ) and*

(iv) $\min_\Delta \max_X |\theta(x)| = \max_X |\theta_0(x)|$.

Furthermore, there is a point $\theta_0 \in \Delta$ satisfying (i) *through* (iv).

This theorem is a special case of Theorem 4.2 proved in Spruill [20] from which it also follows, as one can easily verify, that the variance of the best linear unbiased estimator of the value of (τ, θ) for the optimal design is $[\max_X |\theta_0(x)|]^{-2}\, \sigma^2/N$.

All of the optimal designs presented here are found in this way, but only in the case of the hollow d-cube of Section 3 is a proof presented. The idea in this and each of the others is simply to solve the problem (iv) and thereby deduce the support points and the correct masses.

Example 1 gives a specific application of Theorem 4.2, which applies to any dimension d, degree $m-1$, and values of r and R, to a case in which $d = 3$.

EXAMPLE 1. If measurements of the values of a multivariate polynomial $\theta(x)$ are contaminated with zero mean random errors and such measurements permitted only within the radii r and R, $0 < r < R < \infty$, then how shall N points of observation be chosen if the average value of θ inside the ball of radius r is to be linearly estimated with the smallest possible variance? An optimal choice is given for a particular case in which the polynomial is a function of $d = 3$ variables, $x = (x_1, x_2, x_3)$, $r = 3/4$, $R = 1$ (this value is used throughout the remainder of the paper since there is no loss of generality in this assumption), θ is of the form (1.2) with $m=3$, and the constants c_α are unknown. The design which places masses (takes the indicated proportion of observations at the indicated points)

$$0.0237 \text{ at the points } (\pm 1, \pm 1, \pm 1)/\sqrt{3},$$
$$0.0211 \text{ at the points } \pm e_j,$$
$$0.0513 \text{ at } (\pm 1, \pm 1, \pm 1)/2,$$
$$0.0456 \text{ at } \pm (3/4)\, e_j$$

is optimal for estimating the average value of θ over the set of $x \in R^3$ such that $\|x\|_2 < \frac{3}{4}$. The optimal variance is $6.41\sigma^2/N$. Here the e_j are the standard basis vectors.

The optimal designs are presented below in three sections. The Appendix contains some details which are necessary in the solution to the problems (iv) and some characterizations of the solutions which enable an efficient computer search when necessary.

2. The Univariate Case

Let $w(x) \geqslant 0$ be an even function, $\int_{-D}^{D} w(x)\, dx = 1$, Θ the space of means be the set $\mathbf{P}_{m-1}$ of univariate polynomials of degree no more than $m-1$, and $X = [-1, -D] \cup [D, 1]$. Here m is at least one. For a set of m distinct points $x_1, \dots, x_m$ denote by

$$\phi_{x_1}, \dots, \phi_{x_m}$$

the Lagrange interpolation polynomials of degree $m-1$ to these points. Let T_k be the kth degree Chebyshev polynomial of the first kind.

Theorem 2.1. *For estimating $\int_{-D}^{D} \theta(x)\, w(x)\, dx$, if $m=1$ then all designs are optimal and if $m=2$ then all symmetric designs are optimal. If $m>2$ is even then the optimal design places masses ξ_j proportional to*

$$\int_{-D}^{D} |\theta_{y_j}(x)|\, w(x)\, dx$$

at the points y_j of

$$I_m = \{-x_1, \dots, -x_{m/2}, x_{m/2}, \dots, x_1\},$$

where

$$x_j = (((1-D^2)/2) \cos[2\pi(j-1)/(m-2)] + (1+D^2)/2)^{1/2}.$$

The optimal variance is

$$V_m = \left(\int_{-D}^{D} T_{(m-2)/2}\left[(1+D^2-2x^2)/(1-D^2)\right] w(x)\,dx\right)^2 (\sigma^2/N).$$

If m is odd then let $v_1, ..., v_{m+1}$ be the points of I_{m+1}. If A is either the set $\{v_1, ..., v_m\}$ or $\{v_2, ..., v_{m+1}\}$ and $\{\phi_s^A(x)\}_{s\in A}$ are the Lagrange interpolation polynomials of degree $m-1$ to the points of A then the design on A assigning masses $\xi^A(s)$ proportional to

$$\int_{-D}^{D} |\phi_s^A(x)|\, w(x)\,dx, \qquad s\in A$$

is optimal. The optimal designs are all convex combinations of these two designs and the optimal variance is V_{m+1}.

The following theorem refers to the optimal estimation of $\theta^{(k)}(0)$ when $X=[-1, -D]\cup[D, 1]$, where $D\in(0, 1)$, $k\in\{0, 1, 2, ..., m-1\}$, and

$$\theta(x)=\sum_{j=0}^{m-1} c_j x^j.$$

In the statement of the theorem we denote by $A_j(x)$ the jth Lagrange interpolation polynomial of degree m to the appropriate $m+1$ points, by $B_j(x)$ the jth of degree $m-2$ to the appropriate $m-1$ points, and by $C_j(x)$ the jth of degree $m-1$ to the appropriate m points. We shall also refer to certain sets, $\mathscr{T}_m$ and $\mathscr{S}_m(D)$ which are defined at the end of the Appendix. Except for the case $\mathscr{S}_m(D)$, with m even and $\sin(\pi/2(m-1))<D$, these are defined in terms of elementary functions.

THEOREM 2.2. *If m is odd and k is even then the design ξ_0 which places masses $\xi_0(x_j)$ proportional to $|A_j^{(k)}(0)|$, $j=1, ..., m+1$, at the $m+1$ points of $\mathscr{T}_m$ is optimal.*

If m is odd and k is odd then the unique optimal design ξ_0 places masses $\xi_0(x_j)$ proportional to $|B_j^{(k)}(0)|$, $j=1, ..., m-1$, at the $m-1$ points of $\mathscr{S}_{m-1}(D)$.

If m is even and k is odd the unique optimal design ξ_0 places masses $\xi_0(x_j)$ proportional to

$$|C_j^{(k)}(0)|, \qquad j=1, ..., m, \tag{2.1}$$

at the m points of $\mathscr{S}_m(D)$.

If m is even and k is even the unique optimal design ξ_0 places masses according to (2.1) at the m points of $\mathscr{T}_{m-1}$.

EXAMPLE 2. To illustrate the application of the theorem, when $m=6$ and $D=0.4$ one can show by first finding the support points that the optimal variances for estimating derivatives are as follows, where the variance is the given factor times σ^2/N.

Derivative	0	1	2	3	4	5
Variance	2.81	5.46	26.30	126.91	272.11	2003.31

Since $\sin(\pi/10)<0.4$ and m is even, the optimal design's support for the odd derivatives is given by the elliptic functions of the Appendix; but to find these quantities it is more convenient to utilize the characterization provided by Theorem 5.2.2 in a numerical search. This search identifies the support points as -1, -0.813, -0.4, 0.4, 0.813, and 1 from which one can easily calculate the weights. Of course, for estimating the kth coefficient the variance is only $1/(k!)^2$ times the stated value.

Note that no claim of unicity of the optimal design is made in the case of m odd and k even. When $k=0$, Spruill [19] gives all optimal designs. They consist of all convex combinations of the two formed by omitting the endpoints -1 and $+1$ from the set of $m+1$ extremal points of q_m^* (see the Appendix below for q_m^*). When $k=m-1$ one can show that a similar situation holds except that all optimal designs are obtained as convex combinations of the two which omit $-D$ and D from the extremal points of q_m^*. The complete characterization of all optimal designs when $k \in \{2, ..., m-3\}$ is unknown.

One can obtain all optimal designs for estimating $\theta^{(k)}(c)$ for c in the complement of X if $k=0$ and for c in the complement of the interior of X if $k>0$. Their supports do not depend upon c. For the other cases, $k \in \{1, ..., m-2\}$, only the special choices of c and X covered above result in known optimal designs. If, for example, $k \in \{1, ..., m-2\}$ and $c \neq 0$ is in the interior of $X=[-1, 1]$ then the optimal designs are unknown except in some isolated special cases covered in Murty and Studden [16].

Extrapolation problems are not of primary concern here but we close this section with a theorem which follows from Theorem A.2.3 and a theorem of Kiefer and Wolfowitz [12]. Let c be in the complement of $[A, E]$ if $k=0$ and in the complement of (A, E) if $k>0$.

THEOREM 2.3. *The optimal design for estimating $\theta^{(k)}(c)$ is supported on the set of points at which $|q_m^*|$ attains its maximum value on X, a set consisting of either m or $m+1$ points, independent of c, and containing A and E. If there are m points $x_1 < \cdots < x_m$ then*

$$\xi_0(x_j)=\frac{|\phi_j^{(k)}(0)|}{\sum_{i=1}^m |\phi_i^{(k)}(0)|},$$

where $\{\phi_j\}_{j=1}^{m}$ are the Lagrange interpolation polynomials of degree $m-1$ to $\{x_1, ..., x_m\}$. If there are $m+1$ points then there is at least one subset of m points for which the design is optimal.

Studden [22] gives an elegant solution to the multivariate extrapolation problem from a convex design space utilizing the one-dimensional results. His technique does not work in our case.

3. The Hollow d-Cube

A proof of the optimal design is given in this one section. Proofs of optimality are not given for the other cases in this paper but use the techniques found here.

Let it be required to estimate the value of

$$(\tau, \theta) = \sum_{i=1}^{d} \left. \frac{\partial^k \theta(x)}{\partial x_i^k} \right|_{x=0} = \sum_{i=1}^{d} \theta_i^{(k)}(0),$$

where $k \in \{0, ..., m-1\}$ is even, θ is given by (1.1),

$$X = \{x \in R^d : D \leqslant \|x\|_\infty \leqslant 1\},$$

and D is in $(0, 1)$. Let $X_0 = [-1, -D] \cup [D, 1]$, p_{km} solve $P_{km-1}(X_0, 0)$ (see the Appendix), and

$$O_m = \{x \in X_0 : |p_{km}(x)| = \|p_{km}\|\}$$

be the oscillation set of p_{km} with $m > 2$. First a preliminary result is needed.

Lemma 3.1. *Let*

$$\theta_0(x) = \left(\sum_{j=1}^{d} \left[\prod_{i \neq j} p_{0m}(x_i) \right] p_{km}(x_j) \right) \Big/ d^2$$

and $S = \{ y e_j : 1 \leqslant j \leqslant d, y \in O_m\}$. Then $\{x \in X : |\theta_0(x)| = \|\theta_0\|\} = S$.

Proof. According to Corollary A.3.1 of the Appendix, if $m-k$ is odd then p_{km} is proportional to q_m^*, while if $m-k$ is even it is proportional to q_{m-1}^*. Furthermore, since k is even, if m is even then p_{km} is proportional to p_{0m-1} and if m is odd then p_{km} is proportional to p_{0m}. Since if m is even we have $p_{0m} = p_{0m-1}$, we have in any case for some $\gamma > 0$,

$$\theta_0(x) = \gamma \prod_{i=1}^{d} p_{0m}(x_i)/d.$$

Since $x \in X$ entails $\max|x_i| \geqslant D$ for at least one coordinate, say x_1, if $x \in X \backslash S$, either there are $d-1$ coordinates 0 and x_1 in $X_0 \backslash O_m$ or x_1 is in X_0 and there is at least one coordinate unequal to 0. In the former case one has

$$|p_{0m}(x_1)| < \|p_{0m}\|$$

and, since for all $x \in [-1, 1]$, $|p_{0m}(x)| \leqslant 1$, we have

$$|\theta_0(x)| < \gamma \|p_{0m}\|/d.$$

In the latter case one has $|p_{0m}(x_1)| \leqslant \|p_{0m}\|$ and for some coordinate $|p_{0m}(x_i)| < 1$. Therefore, $|\theta_0(x)| < \gamma\|p_{0m}\|/d$. When $x \in S$,

$$|\theta_0(x)| = \gamma \|p_{0m}\|/d. \quad \blacksquare$$

Since k is even, the only oscillation sets O_m which arise are elementary. They are $\mathscr{T}_m$ for m odd and $\mathscr{T}_{m-1}$ for m even, where these sets are defined at the end of the Appendix.

THEOREM 3.1. *For estimating $\sum_{i=1}^d \theta_i^{(k)}(0)$, $k \geqslant 0$ even, if $m=1$ then all designs are optimal. If $m=2$ then all designs ξ which are invariant under the group of reflections through the origin are optimal. If $m>2$ is even then the design which places masses $\xi(x_i e_j)$ proportional to $|\phi_{x_i}^{(k)}(0)|$, for $i=1, ..., m$ and $j=1, ..., d$ where $\{x_i\}_{i=1}^m = \mathscr{T}_{m-1}$ is optimal with variance $d^2\|p_{km}\|^{-2}\sigma^2/N$. If $m>2$ is odd then the design which places masses $\xi(x_i e_j)$ proportional to $|\phi_{x_i}^{(k)}(0)|$, for $i=1, ..., m+1$ and $j=1, ..., d$, $\{x_i\}_{i=1}^{m+1} = \mathscr{T}_m$, is optimal with variance $d^2\|p_{km}\|^{-2}\sigma^2/N$.*

Proof. The cases $m=1$ and $m=2$ are easy. We verify the case $m>2$ odd. Let

$$\theta_0(x) = \sum_{j=1}^d \left[\prod_{i \neq j} p_{0m}(x_i)\right] p_{km}(x_j)/d^2.$$

Then $(\tau, \theta_0) = 1$ and from the Lemma 3.1 we know that (i) of Theorem 1.1 holds.

To verify (ii) it suffices by linearity to verify it for

$$x^\alpha = \prod_{i=1}^d x_i^{\alpha_i}$$

and $\|\alpha\|_\infty \leqslant m-1$. One has in this case

$$\int x^\alpha \theta_0(x)\, d\xi(x) = \sum_{y \in \mathscr{T}_m} \sum_{i=1}^d (y e_i)^\alpha\, \theta_0(y e_i)\, \xi(y e_i), \tag{3.1}$$

which is zero if α is not of the form ue_v for some $u \in \{0, ..., m-1\}$, $v \in \{1, ..., d\}$. For α of this form the expression (3.1) is

$$d\gamma \sum_{y \in \mathcal{T}_m} y^u [p_{0m}(0)]^{d-1} p_{km}(y)) \frac{|\phi_y^{(k)}(0)|}{d \sum_{z \in \mathcal{T}_m} |\phi_z^{(k)}(0)|}$$

$$= \frac{\gamma \| p_{km} \|}{\sum_{y \in \mathcal{T}_m} |\phi_y^{(k)}(0)|} \sum_{y \in \mathcal{T}_m} y^u \phi_y^{(k)}(0))$$

$$= \frac{\gamma \| p_{km} \|}{\sum_{y \in \mathcal{T}_m} |\phi_y^{(k)}(0)|} \zeta^{(k)}(0),$$

where $\zeta(x) = x^u$. Therefore $\int x^\alpha \theta_0(x) \, d\xi(x)) = \gamma \| p_{km} \|^2 k!$ if $\alpha = ke_v$, $v \in \{1, ..., d\}$, and zero otherwise.

To verify (iii) suppose that for some θ one has $\int \theta^2(x) \, d\xi(x)) = 0$. Let

$$\theta^*(x) = \frac{1}{|G|} \sum_{g \in G} \theta(gx),$$

where G is the group of permutations and coordinate sign changes on R^d. Then by Jensen's inequality

$$\int \theta^{*2}(x) \, d\xi(x) \leqslant \frac{1}{|G|} \sum_{g \in G} \int \theta^2(gx) \, d\xi(x)$$

and, since ξ is invariant under G, one has

$$\int \theta^{*2}(x) \, d\xi(x) = 0. \tag{3.2}$$

For $y \in \mathcal{T}_m$ and all $v \in \{1, ..., d\}$,

$$\theta^*(ye_v) = \frac{1}{d} \sum_{i=0}^{(m-1)/2} \sum_{v=1}^{d} c_{2ie_v} y^{2i},$$

so that (3.2) entails a polynomial of degree $m-1$ vanishing at $m+1$ points. This implies that $(\tau, \theta^*) = \sum_{v=1}^d c_{ke_v} = 0$. Because $(\tau, \theta^*) = (\tau, \theta)$, (iii) has been verified.

The variance expression follows from $\theta_0(x) = \gamma A d \prod_{i=1}^d p_{0m}(x_i)$, since $(\tau, \theta_0) = \gamma A d^2 p_{0m}^{(k)}(0) = \gamma d^2 = 1$. Thus the variance is

$$V = \frac{\sigma^2}{N\gamma \| p_{km} \|^2} = d^2 \| p_{km} \|^{-2} \sigma^2 / N. \quad \blacksquare$$

Example 3. Employing the figures from Example 2, Theorem 3.1 shows that if $m = 6$, $d = 3$, and $D = 0.4$ then the variance of the optimal estimate of the Laplacian, $\Delta\theta$, evaluated at 0 is $9(26.30)\,\sigma^2/N$.

4. The Hollow d-Sphere

Here the design space is assumed to be

$$X = \{x \in R^d : D \leqslant \|x\|_2 \leqslant 1\},$$

with the mean function given in (1.2) and $D \in (0, 1)$.

The designs reported in this section are of two types. The first are designs which, like those above in Sections 2 and 3, are on a finite number of points and easily implemented. The second are designs which will be expressed as uniform masses on a finite number of spheres. The former category arises in estimation of

$$\sum_{i=1}^{d} \theta_i^{(k)}(0),$$

where $k \geqslant 0$ is even, while the latter arises in the estimation of

$$\int_{\|x\| < D} \theta(x)\, w(\|x\|_2)\, dV.$$

Designs of the latter type also arise as D-optimal designs on spherical regions. See Kiefer [11]. To implement such designs one replaces the uniform masses with discrete masses so as to preserve the design matrix as is illustrated in Farrell, Kiefer, and Walbran [6].

Theorem 4.1. *For estimating* $\sum_{i=1}^{d} \theta_i^{(k)}(0)$, $0 \leqslant k \leqslant m-1$, k *even, the designs of Theorem* 3.1 *are optimal and the variances are the same.*

The proof proceeds as in Theorem 3.1 but with

$$\theta_0(x) = \gamma p_{km}(\|x\|_2).$$

Let w be a probability measure on $\|x\|_2 < D$ invariant under the orthogonal group.

Theorem 4.2. *For estimating*

$$\int_{\|x\|_2 < D} \theta(x)\, dw(x)$$

if $m = 1$ *all designs are optimal and if* $m = 2$ *then any design invariant under the group of reflections through the origin is optimal. If* $m > 2$ *is even then the design* ξ_0 *which places masses uniformly on the each of the* $m/2$ *shells of radii*

$$r_j = (((1 - D^2)/2)\cos[2\pi(j-1)/(m-2)] + (1 + D^2)/2)^{1/2}$$

for $j=1$ to $m/2$ assigning masses ξ_j proportional to

$$\int_{\|x\|<D} |\phi_{r_j^2}(\|x\|^2)|\, dw(x)$$

to the jth *shell is optimal. The optimal variance is*

$$V_m = \left(\int_{\|x\|<D} p_{0m}(\|x\|)\, dw(x)\right)^2 \frac{\sigma^2}{N\|p_{0m}\|^2}.$$

If m is odd then the design as above on the $(m+1)/2$ radii

$$r_j = (((1-D^2)/2)\cos[2\pi(j-1)/(m-1)] + (1+D^2)/2)^{1/2}$$

for $j=1$ to $(m+1)/2$ is optimal with variance V_{m+1}.

The proof proceeds as the others using here

$$\theta_0(x) = \left(\int_{\|x\|<D} p_{0m}(\|x\|)\, dw(x)\right)^{-1} p_{0m}(\|x\|).$$

One can also simplify the calculation of the masses. The measure w is the cross product of a measure v on $[0, D)$ and uniform measure on the unit sphere in R^d. The masses are therefore proportional to the quantities, when m is even,

$$\int_0^D |\phi_{r_j^2}(t^2)|\, dv(t)$$

for $j=1$ to $m/2$, where $\{\phi_{r_i^2}\}_{i=1}^{m/2}$ are the Lagrange interpolation polynomials of degree $m/2-1$ to the points r_i^2. It is clear that when v is concentrated at a single radius $r \in (0, D)$ the designs given above are optimal for estimating the value of a surface integral of θ over this ball inside the hollow sphere.

One example of the application of Theorem 4.2 was provided in the Introduction; here is another.

EXAMPLE 4. Under the same assumptions as in Example 1 of the Introduction, for estimating the surface integral of θ over the ball of radius $\frac{1}{4}$ the design which places masses

$$0.0261 \text{ at the points } (\pm 1, \pm 1, \pm 1)/\sqrt{3},$$

$$0.0232 \text{ at the points } \pm e_j,$$

$$0.0489 \text{ at } (\pm 1, \pm 1, \pm 1)/2,$$

$$0.0435 \text{ at } \pm(3/4)\, e_j$$

is optimal with an associated variance of $\pi^2\sigma^2/N$.

APPENDIX

A.1. *Introduction*

Deboor and Rice [5] developed a Remes exchange algorithm for solving the problems $P_{0m-1}(X, 0)$. For arbitrary $c \in (-\infty, \infty)$ the problems are

$P_{0m-1}(X, c)$: find $p \in \mathbf{P}_{m-1}$ satisfying $p(c) = 1$ with the minimal value of $\|p\|$.

Here $\mathbf{P}_{m-1}$ is the collection of polynomials of degree $m-1$ or less, and

$$\|p\| = \max_X |p(x)|.$$

They considered $X = [A, B] \cup [D, E]$ with $B < 0 < D$. Like Deboor and Rice, Lebedev [15] was motivated in his studies by the use of the solutions to provide good parameter choices in a Richardson iteration for indefinite linear systems. He studied the above problem for X a union of special disjoint intervals, none containing the point 0, and provided in the case of two symmetric disjoint intervals $X = [A, B] \cup [-B, -A]$ explicit solutions to $P_{0m-1}(X, 0)$ for odd m.

Attlestam [4] obtained convergence estimates for solutions of definite linear systems using the conjugate gradient method by studying Chebyshev polynomials of disjoint intervals, $X = [A, B] \cup [D, E]$ when $0 < A$ or $E < 0$. Introducing the problems

$MP_{m-1}(X)$: find p, the monic polynomial in $\mathbf{P}_{m-1}$ minimizing $\|p\|$,

Attlestam used the fact that if $0 < A$ or $E < 0$ and q_m^* solves $MP_{m-1}(X)$ then $q_m^*(x)/q_m^*(0)$ solves $P_{0m-1}(X, 0)$. She also noted that if $B < 0 < D$ then in general the solutions to $P_{0m-1}(X, 0)$ and $MP_{m-1}(X)$ are unrelated.

Achieser [2] had provided explicit solutions to $MP_{m-1}(X)$ when $X = [-1, -D] \cup [D, 1]$ and asymptotic formulas for the minimal deviation when $X = [-1, B] \cup [D, 1]$. One can verify that Lebedev's solution to $P_{0m-1}(X, 0)$ and Achieser's solution to $MP_{m-1}(X)$ are proportional when m is odd and $X = [-1, -D] \cup [D, 1]$.

We show that in general the solutions can be proportional only if m is odd. Furthermore, for the problems

$P_{km-1}(X, c)$: find $p \in \mathbf{P}_{m-1}$ satisfying $p^{(k)}(c) = 1$ with the minimal value of $\|p\|$

it is shown that if $m-k$ is odd then the solution P_{km}, when $B - A = E - D$ and $c = 0$, is proportional to q_m^* and if $m-k$ is even is proportional to q_{m-1}^*. Achieser [2] stated this for m even and k odd.

A.2. *Characterizations*

For an arbitrary $X = [A, B] \cup [D, E]$, $-\infty < A < B < D < E < +\infty$, and continuous function f on X introduce the problems

$$A_{m-1}(f)\text{: minimize } \|f - p\| \text{ over } p \in \mathbf{P}_{m-1}.$$

It is assumed henceforth that $m > 2$. The next result follows from the usual alternation theorem (see Cheney [5a]).

LEMMA A.2.1. *There is a unique solution p^* to $A_{m-1}(f)$ and there are subsets R of $[A, B]$ and S of $[D, E]$ whose union contains at least $m+1$ points, numbers q_R and q_S in $\{0, 1\}$, points $x_1 < \cdots < x_L$ in R and points $x_{L+1} < \cdots < x_{m+1}$ in S such that*

$$f(x_j) - p^*(x_j) = \begin{cases} (-1)^{j-q_R}\|f - p^*\|, & 1 \leqslant j \leqslant L, \\ (-1)^{j-q_S}\|f - p^*\|, & L+1 \leqslant j \leqslant m+1. \end{cases} \tag{5.2.1}$$

The solution q_m^* to $MP_{m-1}(X)$ *is* $q_m^*(x) = x^{m-1} - v^*(x)$, where $v^*(x)$ solves $A_{m-2}(x^{m-1})$ so q_m^* oscillates properly; that is, in the form described by (5.2.1), in at least $m\text{-}2 + 2 = m$ points.

The next theorem is proved in Spruill [18]. Also see DeBoor and Rice [5].

THEOREM A.2.1. *There is a unique solution $p_m \in \mathbf{P}_{m-1}$ to $P_{0m-1}(X, c)$. Furthermore, p is the solution if and only if p oscillates properly on X in at least m points, $p(c) = 1$, and has B and D among the oscillation points with $p(B) = p(D) = \|p\|$.*

One can show as a corollary that all solutions to $P_{0m-1}(X, c)$, $c \in (B, D)$ are proportional.

It is not difficult to show that if q_m^* solves $MP_{m-1}(X)$ and oscillates properly in exactly m points then $q_m^*(x_L) = -q_m^*(x_{L+1})$ and that this also entails both A and E in the oscillation set; nor is it difficult to show that if a monic polynomial $q \in \mathbf{P}_{m-1}$ oscillates properly in $m+1$ points then m is odd, $q(x_L) = q(x_{L+1})$, and B and D are in the oscillation set.

The next theorem is easily proven using these facts.

THEOREM A.2.2. *The monic polynomial q^* of degree $m-1$ solves $MP_{m-1}(X)$ if either*

(a) *q^* oscillates properly at $m+1$ points or*

(b) *q^* oscillates properly at m points, A and E are in the oscillation set, and $q^*(x_L) = -q^*(x_{L+1})$.*

Conversely, if q_m^* solves $MP_{m-1}(X)$ then q_m^* satisfies either (a) *or* (*b*).

The following theorem now follows easily.

THEOREM A.2.3. *Let* q_m^* *solve* $MP_{m-1}(X)$ *and* c *be in the complement of* $[A, E]$ *if* $k \geqslant 1$ *or in the complement of* (A, E) *if* $k = 0$. *Let* p *minimize* $\|p\|$ *among polynomials in* $\mathbf{P}_{m-1}$ *for which* $p^{(k)}(c) = 1$. *If* $0 \leqslant k \leqslant m-1$ *then* $q_m^{*(k)}(c) \neq 0$ *and* $p(x) = q_m^*(x)/q_m^{*(k)}(c)$.

A.3. *The Symmetric Case*

If $B - A = E - D$ one can assume without loss of generality that $X = [-1, -D] \cup [D, 1]$. In this case it is easy to see that $p_m = p_{m+1}$ for $m = 2n - 1$ and that the degree of p_{m+1} is $2n - 2$, where p_m solves $P_{0m-1}(X, 0)$. This shows that p_m is proportional to q_m^*. More is true.

Fix $k \in \{0, ..., m-1\}$ and consider the problem

$$MP_{km-1}(X)\colon \min_{a_i \in R\, i \neq k} \max_{x \in X} \left| x^k - \sum_{\substack{i=0 \\ i \neq k}}^{m-1} a_i x^i \right|.$$

The minimizing polynomial of degree $\leqslant m-1$ shall be known as the solution. Let $q_m^*(x) = \sum_{i=0}^{m-1} A_{im} x^i$ solve $MP_{m-1}(X) = MP_{m-1m-1}(X)$. The proof of the following theorem can be carried out along the lines of the proof of W. A. Markov's theorem (see Natanson [17]).

THEOREM A.3.1. *If* $m - k$ *is odd then the solution to* $MP_{km-1}(X)$ *is* q_m^*/A_{km} *and if* $m - k$ *is even is* q_{m-1}^*/A_{km-1}.

COROLLARY A.3.1. *If* $m - k$ *is odd then the solution to* $P_{km-1}(X, 0)$ *is proportional to* q_m^* *and if* $m - k$ *is even it is proportional to* q_{m-1}^*.

To describe the oscillation set of q_m^* some preliminary facts are required. According to Achieser [2] or Achieser [3, in Section 27 of Addenda and Problems], if m is odd the solution q_m^* is proportional to $T_{(m-1)/2}[(1 + D^2 - 2x^2)/(1 - D^2)]$ and when m is even, to the function F given in (5.3.1) when

$$x = \frac{Dcn(i\rho K')}{\sqrt{D^2 - sn^2(i\rho K')}}.$$

Suppose that m is even and $\sin[\pi/2(m-1)] < D$. Then there is a unique number $\eta \in (0, 1)$, called the modulus of the elliptic function, which satisfies

$$\int_0^{\sin^{-1} D} \frac{d\theta}{\sqrt{1 - \eta \sin^2 \theta}} \bigg/ \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - \eta \sin^2 \theta}} = \frac{1}{m-1}.$$

Set

$$K = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - \eta \sin^2 \theta}}$$

and let iK' be determined as usual (see Abramowitz and Stegun [1] for details). Let

$$v_1(z) = 2 \sum_{j \geq 0} (-1)^j q^{(j+1/2)^2} \sin[(2j+1)z],$$

where $q = e^{-\pi K'/K}$ and define, for $\rho \in [0, 1]$,

$$F(\rho) = \cos(\theta(\rho)), \tag{5.3.1}$$

where

$$\theta(\rho) = \left(\frac{m-1}{2}\right) \arg\left[\frac{v_1(\pi/2(m-1) + i\rho\pi K'/2K)}{v_1(\pi/2(m-1) - i\rho\pi K'/2K)}\right].$$

One can verify that $\theta(0) = 0$ and $\theta(1) = -(m/2 - 1)\pi$. Consequently there are values $\rho_j \in [0, 1]$, $j = 1, ..., m/2$ for which

$$\frac{v_1(\pi/2(m-1) + i\rho\pi K'/2K)}{v_1(\pi/2(m-1) - i\rho\pi K'/2K)} = e^{-i2(j-1)\pi/(m-1)}.$$

Let

$$y_j = \frac{Dcn(i\rho_j K')}{\sqrt{D^2 - sn^2(i\rho_j K')}}$$

for $j = 1$ to $m/2$ and $\mathscr{S}_m(D)$ consist of the m points $\{\pm y_j\}_{j=1}^{m/2}$. If $\sin[\pi/2(m-1)] \geq D$ let $x_j = -\cos[(j-1)\pi/(m-1)]$ and $\mathscr{S}_m(D) = \{x_j\}_{j=1}^{m}$.

When m is odd let

$$y_j = ((1 - D^2/2)\cos[2\pi(j-1)/(m-1)] + (1 + D^2)/2)^{1/2}$$

and $\mathscr{T}_m$ consist of the $m+1$ points $\{\pm y_j\}_{j=1}^{(m+1)/2}$.

References

[1] Abramowitz, M., and Stegun, I. (1970). *Handbook of Mathematical Functions.* Dover, New York.

[2] Achieser, N. I. (1928). Uber einigen functionen die in gegebenen intervallen am wenigsten von null abwiechen. *Izv. Kazansk. Fiz. Mat. Obshch.*

[3] Achieser, N. I. (1956). *Theory of Approximation.* Ungar, New York.

[4] ATTLESTAM, B. (1977). *Tchebycheff Polynomials for Sets Consisting of Two Disjoint Intervals with Application to Conergence Estimates for the Conjugate Gradient Method.* Res. Rep. 77.06 R, Dept. Computer Sciences, Chalmers University of Technology and the University of Goteborg, Sweden.

[5a] CHENEY, E. W. (1966). *Introduction to Approximation Theory.* McGraw–Hill, New York.

[5] DEBOOR, C., AND RICE, J. (1982). Extremal polynomials with application to Richardson iteration for indefinite linear systems, *SIAM J. Sci. Comput.* 3 47–57.

[6] FARRELL, R., KIEFER, J., AND WALBRAN, A. (1967). Optimum multivariate designs. *Proceedings, Fifth Berkeley Symposium Math. Statist. and Probability*, pp. 113–138.

[7] HERZBERG, A., AND COX, D. (1977). Some optimal designs for interpolation and extrapolation. *Biometrika* **59** 551–561.

[8] HOEL, P. (1965a). Minimax designs in two dimensional regression. *Ann. Math. Statist.* **36** 1097–1106.

[9] HOEL, P. (1965b). Optimum designs for polynomial extrapolation. *Ann. Math. Statist.* **36** 1483–1493.

[10] KARLIN, S., AND STUDDEN, W. (1966). Optimal experimental designs. *Ann. Math. Statist.* **37** 783–815.

[11] KIEFER, J. (1960). Optimum experimental designs *V* with applications to systematic and rotatable designs. In *Proceedings Fourth Berkeley Symp. Math. Statistics and Probability*, pp. 381–405.

[12] KIEFER, J., AND WOLFOWITZ, J. (1959). Optimum designs in regression problems. *Ann. Math. Statist.* **30** 271–294.

[13] KIEFER, J., AND WOLFOWITZ, J. (1964a). Optimum extrapolation and interpolation designs I. *Ann. Inst. Statist. Math.* **16** 79–108.

[14] KIEFER, J., AND WOLFOWITZ, J. (1964b). Optimum extrapolation and interpolation designs II. *Ann. Inst. Statist. Math.* **16** 295–303.

[15] LEBEDEV, V. (1969). Iterative methods for solving operator equations with a spectrum contained in several intervals. *U.S.S.R. Comput. Math. Phys.* **9** 17–24.

[16] MURTY, V., AND STUDDEN, W. (1972). Optimal designs for estimating the slope of a polynomial regression. *J. Amer. Statist. Assoc.* **67** 869–873.

[17] NATANSON, I. (1964). *Constructive Function Theory 1.* Ungar, New York.

[18] SPRUILL, M. (1986). *Optimal Experimental Designs.* Final Technical Report N.S.F. Grant DMS 8401759.

[19] SPRUILL, M. (1987a). Optimal designs for interpolation. *J. Statist. Plann. Infer.* **16** 219–229.

[20] SPRUILL, M. (1987b). Optimal extrapolation of derivatives. *Metrika.* **34** 45–60.

[21] STUDDEN, W. (1968). Optimal designs on Tchebycheff points. *Ann. Math. Statist.* **39** 1435–1447.

[22] STUDDEN, W. (1971). Optimal designs for multivariate polynomial extrapolation. *Ann. Math. Statist.* **42** 828–832.